

Three-dimensional structure of *Schistosoma japonicum* glutathione S-transferase fused with a six-amino acid conserved neutralizing epitope of gp41 from HIV



KAP LIM,¹ JOSEPH X. HO,¹ KIM KEELING,¹ GARY L. GILLILAND,² XINHUA JI,²
FLORIAN RÜKER,³ AND DANIEL C. CARTER¹

¹ ES 76 Biophysics Branch, Laboratory for Structural Biology, George C. Marshall Space Flight Center, National Aeronautics and Space Administration, Huntsville, Alabama 35812

² Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, and National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, Maryland 20850

³ Institut für Angewandte Mikrobiologie, Universität für Bodenkultur, A-1190 Wien, Austria

(RECEIVED June 17, 1994; ACCEPTED September 5, 1994)

Abstract

The 3-dimensional crystal structure of glutathione S-transferase (GST) of *Schistosoma japonicum* (Sj) fused with a conserved neutralizing epitope on gp41 (glycoprotein, 41 kDa) of human immunodeficiency virus type 1 (HIV-1) (Muster T et al., 1993, *J Virol* 67:6642–6647) was determined at 2.5 Å resolution. The structure of the 3-3 isozyme rat GST of the μ gene class (Ji X, Zhang P, Armstrong RN, Gilliland GL, 1992, *Biochemistry* 31:10169–10184) was used as a molecular replacement model. The structure consists of a 4-stranded β -sheet and 3 α -helices in domain 1 and 5 α -helices in domain 2. The space group of the Sj GST crystal is P4₃2₁2, with unit cell dimensions of $a = b = 94.7$ Å, and $c = 58.1$ Å. The crystal has 1 GST monomer per asymmetric unit, and 2 monomers that form an active dimer are related by crystallographic 2-fold symmetry. In the binding site, the ordered structure of reduced glutathione is observed. The gp41 peptide (Glu-Leu-Asp-Lys-Trp-Ala) fused to the C-terminus of Sj GST forms a loop stabilized by symmetry-related GSTs. The Sj GST structure is compared with previously determined GST structures of mammalian gene classes μ , α , and π . Conserved amino acid residues among the 4 GSTs that are important for hydrophobic and hydrophilic interactions for dimer association and glutathione binding are discussed.

Keywords: fusion protein; glutathione S-transferase; human immunodeficiency virus type 1; *Schistosoma japonicum*; 3-dimensional structure

Glutathione S-transferases (GSTs) are a family of multifunctional enzymes that catalyze the nucleophilic addition of the reduced thiol of glutathione to a variety of electrophiles (Mannervik & Danielson, 1988; Armstrong, 1991; Gilliland, 1993; Rushmore & Pickett, 1993). They are important enzymes involved in the metabolism of potentially toxic alkylating agents. Cytosolic GSTs, dimeric proteins with a molecular mass of each subunit of about 26 kDa, are also involved in intracellular binding and transport of hydrophobic ligands and participate in the synthesis of prostaglandins and leukotrienes. Specific molecular forms of GST are known to be expressed in the cellular mechanisms of drug resistance (Tsuchida & Sato, 1992).

GSTs from a variety of organisms have been identified and characterized (Hiratsuka et al., 1990; Dominey et al., 1991; Meyer et al., 1991; Mannervik et al., 1992; Trottein et al., 1992; Mignogna et al., 1993), including the detection in a wide range of parasitic helminths (Brophy & Barrett, 1990; Brophy et al., 1990). Cytosolic GSTs of the genus *Schistosoma* participate in the immunogenicity to the vertebrate host and have been suggested as potential components of a vaccine against schistosomiasis (Capron et al., 1987; Sher et al., 1989). In addition, the development of a gene expression system employing the GST from *Schistosoma japonicum* (Sj) allows the fusion of a protein or protein fragment to the C-terminus of the enzyme (Smith & Johnson, 1988; Fainsod et al., 1991; Menéndez-Arias et al., 1992; Oshima-Hirayama et al., 1993; Slany et al., 1993). The fusion peptide or protein is coupled to GST through a linker that can be cleaved by a specific protease, thus providing a useful sys-

Reprint requests to: Daniel C. Carter, ES 76 Biophysics Branch, George C. Marshall Space Flight Center, NASA, Huntsville, Alabama 35812; e-mail: carter@lsb.msfc.nasa.gov.

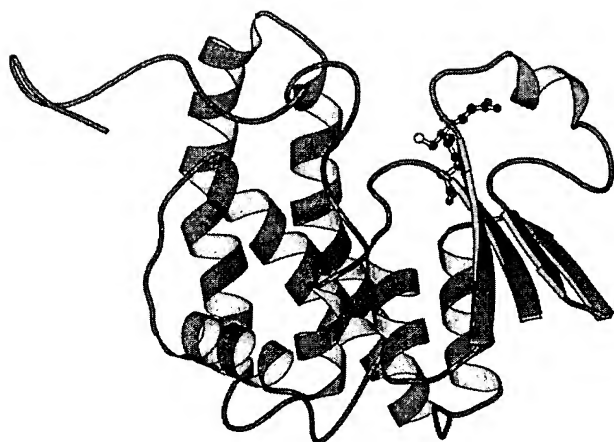


Fig. 1. *Sj* GST with fused peptide in a ribbon diagram. This figure was produced with the MOLSCRIPT program (Kraulis, 1991).

tem for cloning, expressing, and purifying the desired protein or peptide product by glutathione-based affinity chromatography.

A number of crystal structures of mammalian GSTs complexed with a variety of ligands in the active site have been recently determined (Reinemer et al., 1991, 1992; Ji et al., 1992; Sinning et al., 1993; García-Sáez et al., 1994). The structures have revealed a common folding pattern with a number of variations in the binding mode of glutathione and especially the xenobiotic substrate binding site. Crystallographic studies have also been instrumental in determining the structural features of the active site essential for catalysis and substrate specificity (Liu et al., 1992, 1993; Manoharan et al., 1992; Ji et al., 1993, 1994).

In this article, the 3-dimensional crystal structure of *Sj* GST with a C-terminal fusion peptide and reduced glutathione in the active site is described and compared to the 3 previously determined GST structures from the mammalian gene classes, α from human (Sinning et al., 1993), μ from rat (Ji et al., 1992), and π from pig (Reinemer et al., 1991). Because *Sj* GST shares a striking sequence homology with the mammalian μ GST (Smith et al., 1986; Hughes, 1993), this structural comparison reveals the shared characteristics of the amino acid residues in the glutathione binding site of both the vertebrate and invertebrate GSTs. The sequence of the fusion peptide is Glu-Leu-Asp-Lys-

Trp-Ala from gp41 (glycoprotein, 41 kDa) of human immunodeficiency virus type 1 (HIV-1). Muster and coworkers (1993) determined that this peptide fused with GST was bound to the human monoclonal antibody (Mab) 2F5. This sequence of gp41 was found to be highly conserved in a variety of HIV-1 isolates, suggesting its potential crossreactivity to the antibodies with properties similar to Mab 2F5. The original goal of this work was to crystallize the complex of *Sj* GST fusion protein with Mab 2F5, but only a crystal of the GST fusion protein was obtained. Nevertheless, because a peptide or protein fragment by itself is often difficult to crystallize, the GST fusion protein represents a great potential for the structure determination of smaller protein fragments. The significance of this fusion method is discussed further elsewhere (Carter et al., 1994).

Results and discussion

Of the 4 GST structures compared, only *Sj* GST has a monomer in its crystal asymmetric unit. For the others, the 2 monomers in a dimer-paired asymmetric unit are called subunits A and B. Following this practice, the *Sj* GST monomer in the asymmetric unit is also designated subunit A, and its symmetry-related counterpart subunit B. Unless specified otherwise, all structural comparisons were made with subunit A of each GST. The atomic coordinates of μ , α , and π GSTs were obtained from Brookhaven Protein Data Bank (identification codes 1GST, 1GUH, and 1GSR, respectively). Furthermore, the amino acid residue names and numbers used in this article are for *Sj* GST; for the others, residue names and numbers are used in conjunction with a specific GST. The coordinates of the *Sj* GST structure are deposited in the Protein Data Bank (identification code 1GNE).

Final model of the structure

The crystal structure of *Sj* GST with the gp41 fusion peptide at 2.5 Å resolution is shown in Figure 1 and Kinemage 1. It is a monomer forming a perfect dimer by crystallographic 2-fold rotation (Fig. 2). It shares similar characteristics with other GSTs: domain 1 (residues 1–84) is a typical GST folding pattern of $\beta\alpha\beta\alpha\beta\beta$, and domain 2 (residues 85–217) consists of 5 α -helices followed by a long stretch of loops (Fig. 3). The summary of least-squares refinement parameters is shown in Table 1. Also, the ϕ - ψ torsion angle plot for main-chain atoms is shown in Fig-



Fig. 2. Stereo view of a C α tracing of *Sj* GST with the fusion peptide in dimer form. Thick and thin lines distinguish the 2 monomers.

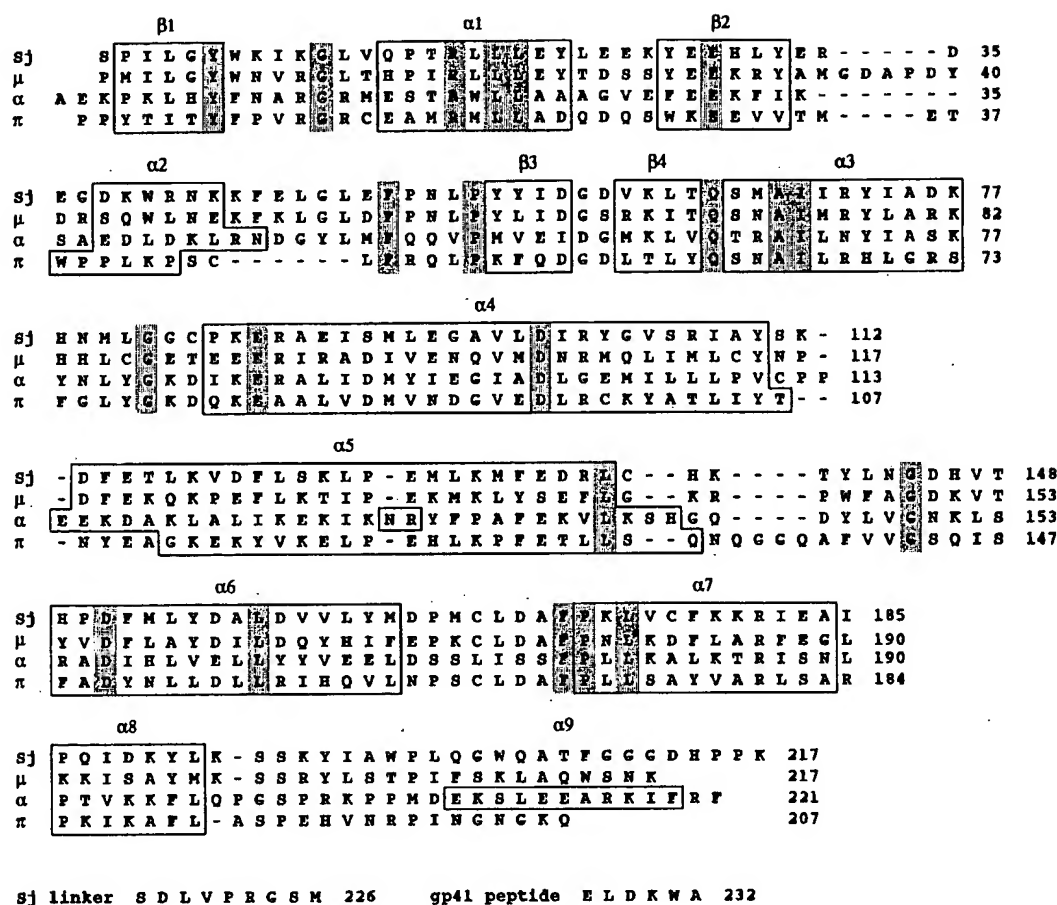


Fig. 3. Amino acid sequences of GSTs. The alignments are based on structural considerations. Secondary structural elements are marked by enclosed lines, and common residues among all GSTs are shaded. The *Sj* GST sequence continues with those of linker peptide and the gp41 fusion peptide.

Table 1. Summary of least-squares refinement parameters of *Sj* GST fusion protein complexed with glutathione

	Target value	Final model
Reflections from 6.0 to 2.3 Å with $I > 1\sigma(I)$		8,917
Reflections from 6.0 to 2.5 Å with $I > 1\sigma(I)$		7,622
Crystallographic R^a factor (6.0–2.3 Å)		0.227
Crystallographic R factor (6.0–2.5 Å)		0.219
RMS deviation from ideal distances (Å)		
Bond distances	0.030	0.020
Angle distances	0.036	0.038
Planar 1–4 distances	0.030	0.023
RMS deviation of bond angle (°)		1.2
RMS deviation from planarity (Å)	0.040	0.034
RMS deviation from ideal chirality (Å ³)	0.200	0.269
Thermal parameter correlations (mean/ $\Delta\beta$)		
Main-chain bend	1.500	1.300
Main-chain angle	2.000	1.950
Side-chain bend	2.000	1.269
Side-chain angle	1.500	1.600

$$^a R = \sum_{hkl} |F_o| - |F_c| / \sum_{hkl} |F_o|.$$

ure 4. The non-glycine residues with near-unfavorable torsion angles are Gln 66, which is involved in glutathione binding, Asn 143, which is in the loop between the α -helices 5 and 6, and Glu 227 in the fusion peptide. There are 2 *cis*-prolines at residues 55 and 201. Furthermore, a highly ordered glutathione structure was determined in the binding site (Fig. 5). The asymmetric unit has 126 water molecules, 30 of which do not have direct interactions with the protein itself but with other water molecules. Most of the water molecules are found in the open region at the dimer interface, around the glutathione binding site, and around the β -strands in domain 1.

Omit maps showed that the residues of the main *Sj* GST structure are well defined. These residues have low temperature factors (4–15 Å²) (Fig. 6). Temperature factors for the fusion peptide are relatively higher (15–30 Å²). Electron density for the linking amino acids Leu 220 and Val 221 is not clearly defined; these are the residues with hydrophobic side chains that are in the bulk solvent region. The electron density map indicates the possibility of alternative conformations for these residues.

The gp41 fusion peptide is extended by the linker peptide (Ser-Asp-Leu-Val-Pro-Arg-Gly-Ser-Met; Fig. 3; see Kinemage 1) to

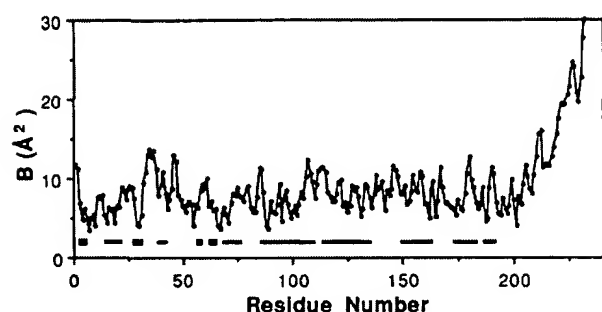


Fig. 6. Plot of temperature factors of main-chain atoms N, C α , and C of *Sj* GST with the fusion peptide. Bars at the bottom of the graph represent the secondary structural elements: thick bars for β -strands and thin bars for α -helices.

perposition of selected C α atoms (Fig. 9). The 4 β -strands and 2 accompanying α -helices 1 and 3 in domain 1 are the most conserved structural features among the 4 GSTs compared here and hence used as the reference structure. The RMS deviation after the first superposition of domain 1 of μ , α , π GSTs to that of *Sj* GST were 0.57 Å, 0.95 Å, and 0.92 Å, respectively. The second superposition was performed to determine the RMS deviations, rotation angles, and magnitude of translation required to superimpose C α atoms of domain 2 (Table 2), which indicate how the tertiary arrangement of domain 2 differs from one GST to another when domain 1 is used as a reference. It is apparent that, although between *Sj*, μ , and π GSTs the magnitude of translation required to move one domain 2 to another for the best fit is about 1 Å, that performed between α GST and the other 3 GSTs is more than 3 Å. This large difference in translation is mostly due to the fact that in α GST, the α -helix 6 (and also α -helices 7 and 8 that follow it) is separated to a much greater extent from the α -helix 4, nearest and parallel to the 2-fold rotational axis for a GST dimer, than in the other GSTs.

The difference in C α atom distances between the α -helices 4 and 6 is about 1.2 Å between α GST and the other GSTs. The rotation angle is also larger by about 2° when α GST is compared with the other 3 GSTs. Furthermore, another source of the structural difference is that the α -helix 4 of *Sj* GST is curved by about 15°, whereas that of μ GST is almost linear.

Hydrophobic and hydrophilic interactions

A common hydrophobic interface between the dimer pair can be observed in all 4 GSTs. The residue in the loop after the α -helix 2 is the conserved Phe 51 and nearby are the residues Leu 64 and Ala 69 that are at the end of the β -strand 4 (Fig. 3). These residues face the residues in subunit B of the dimer pair, Ala 89, Met 93, Leu 94, and Phe 132. The sequence number and identity of residues that participate in this hydrophobic interaction differ slightly in the other GSTs, but the above-mentioned residues make up the core structure for the hydrophobic interaction between the GST dimer pairs. The region along the dimer 2-fold axis is a relatively open hydrophilic channel, and flanking this channel on both sides of the dimer is this hydrophobic interface acting as a stabilizing force for the GST dimer.

The hydrophilic interactions between the dimer pair is more variable than the hydrophobic interactions. Still, certain key sequence positions can be distinguished: Gln 66, Arg 72, Asp 76, Arg 88, Ser 92, and Asp 100. These residues are located in the α -helices 3 and 4, and together with the corresponding α -helices of subunit B, they are positioned nearly parallel along the 2-fold axis and form a hydrophilic core. The residue Gln 66 participates in the glutathione binding and interacts with Asp 100 of subunit B. The residue Arg 72 interacts with its own symmetry-related residue in subunit B. The residue Asp 76 interacts with Arg 88 of subunit B. In μ GST, the polarity is changed so that Arg 81 interacts with Glu 90 and Asp 97 of subunit B. In π GST, Arg 77 interacts with Asp 88 of subunit B. The residue Ser 92 shows no direct interaction with subunit B in *Sj* GST; however, in the

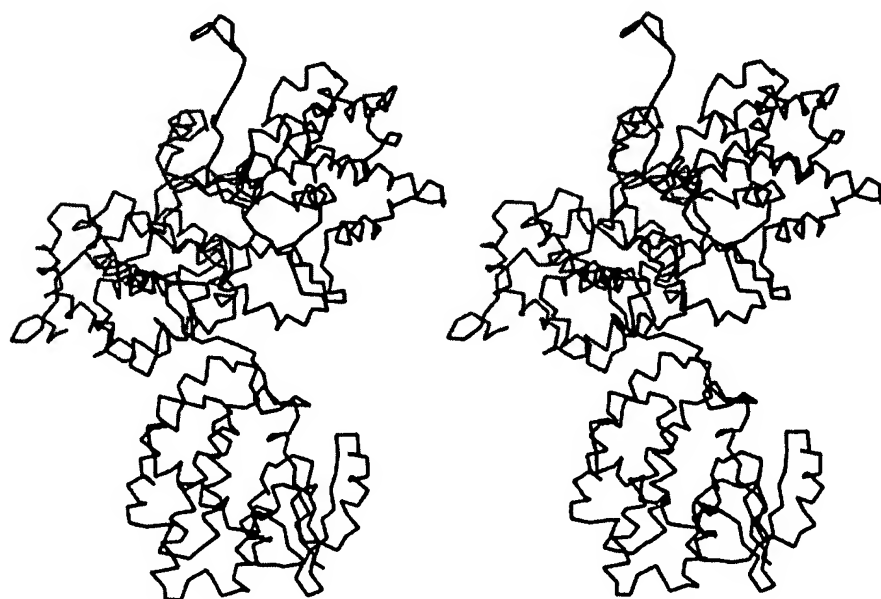


Fig. 7. Stereo view of a C α tracing of *Sj* GST with its fusion peptide interacting with symmetry-related GSTs.

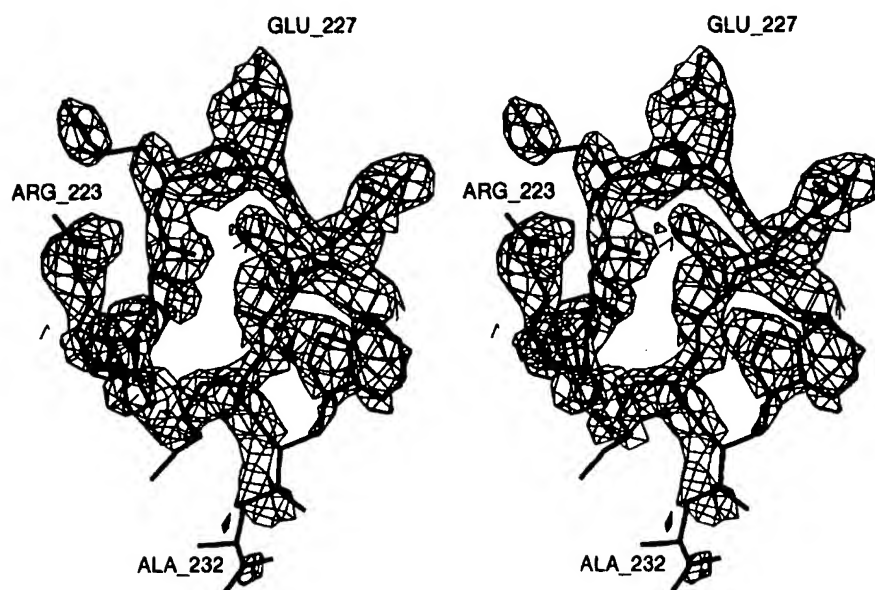


Fig. 8. Stereo view of $F_o - F_c$ electron density map of the gp41 fusion peptide contoured at the 2.5σ level.

other 3 GSTs, this sequence position is aspartate, which shows interactions with a positively charged residue in subunit B.

There are other hydrophilic residues that are conserved in all 4 GSTs at key positions in the sequence. One of these is the residue Asp 151, which is at the beginning portion of the α -helix 6. Its side-chain atoms are observed to interact with the main-chain N atoms of Leu 142 and Thr 148, which are in the loop region after the long α -helix 5. It would seem that the role of Asp 151 is to stabilize the loop region between 2 α -helices. Indeed, a site-directed mutagenesis study showed that the aspartate at this position in π GST is important for thermostability of the protein (Kong et al., 1993). Another amino acid with a similar configuration is Glu 87; it is at the beginning portion of α -helix 4, and its side chain interacts with the N atom of Gly 83, which is in the 78–84 loop connecting the 2 domains of GST. Furthermore, aspartate residues at the β -turn in domain 1 are observed to make interesting interactions with other residues. The side chains of Asp 59 and Lys 77 interact with each other, and Asp 61 is close to Lys 86 of subunit B.

The C-terminal region of *Sj* GST (residues 213–217) essentially sits on the surface of domain 2. It is also close to the 33–35 loop, a shortened version of the 33–42 loop of μ GST, that links the β -strand 2 and α -helix 2 in domain 1. The amino acid residues of these 2 regions interact closely with each other. In *Sj* GST, the side chain of Glu 33 forms a hydrogen bond with the ring atom N ϵ 1 of Trp 205; in μ GST, the analogous interaction is between Asp 36 and Lys 210, though in this case Asp 36 is 2 residues farther down both in sequence and spatial relationship than Glu 33 of *Sj* GST in the interaction with domain 2. The next residue in *Sj* GST is Arg 34, and its side chain interacts with the side chain of Asp 213 and the carbonyl O atom of Gly 212. The corresponding residue in μ GST is Met 34, and rather than its side chain pointing outward as that of Arg 34 of *Sj* GST, it faces into the protein core. The 33–42 loop of μ GST interacts with its own symmetry-related loop and residues of domain 2. In π GST, although the corresponding loops have similar conformations, they are farther apart than in *Sj* GST. This allows the

symmetry-related GST to move into this gap. In α GST, the C-terminal α -helix 9 points toward the glutathione binding site and runs nearly parallel with the α -helix 2. There are hydrophobic interactions between Ile 218, Phe 219, Val 54, Leu 40, and Ile 34. The hydrophilic interaction between Arg 220 and Asp 41 is also observed. Unlike μ GST and π GST, α GST does not have a close symmetry-related interaction at this junction; it would seem that the α -helix 9 is sufficient to close the gap between the 2 domains of GST. Thus, either the close interactions between the 2 domains or crystal packing at this region encloses the glutathione binding site.

Glutathione binding site

In all 4 GSTs, an extended form of glutathione is bound to domain 1. Generally, there are 4 types of interaction with glutathione in the binding site (Figs. 5B, 10): (1) stabilization and orientation of the γ -glutamate of glutathione, (2) alignment of the glutathione peptide backbone, (3) stabilization of the terminal carboxylate of glycine, and (4) interaction with the sulfhydryl of cysteine for enzymatic activation. The binding site residues that are common in all 4 GSTs are Tyr 6, Asn 53, Leu 54, Gln 66, and Ser 67 in subunit A, and Asp 100 of subunit B (Fig. 10A; Kinemages 1 and 2).

Association and orientation of the γ -glutamate is accomplished by either a Ser-Gln or Thr-Gln sequence that form hydrogen bonds with the carboxyl group (atoms O α 1 and O α 2). With the exception of π GST, this configuration is further stabilized by a salt bridge between the atom N of γ -glutamate and Asp 100 of subunit B. In π GST, the γ -glutamyl carboxylate group is rotated approximately 180° primarily by a more favorable salt bridge with the side chain of Arg 13. The arginine residue at this sequence location is also present in α GST. In contrast, leucine is the residue in *Sj* GST and μ GST (Fig. 3).

Hydrogen bonds to the glutathione peptide backbone are produced by another important dipeptide sequence, Leu-Asn in *Sj* GST and μ GST, Val-Gln in α GST, and Leu-Gln in π GST, and

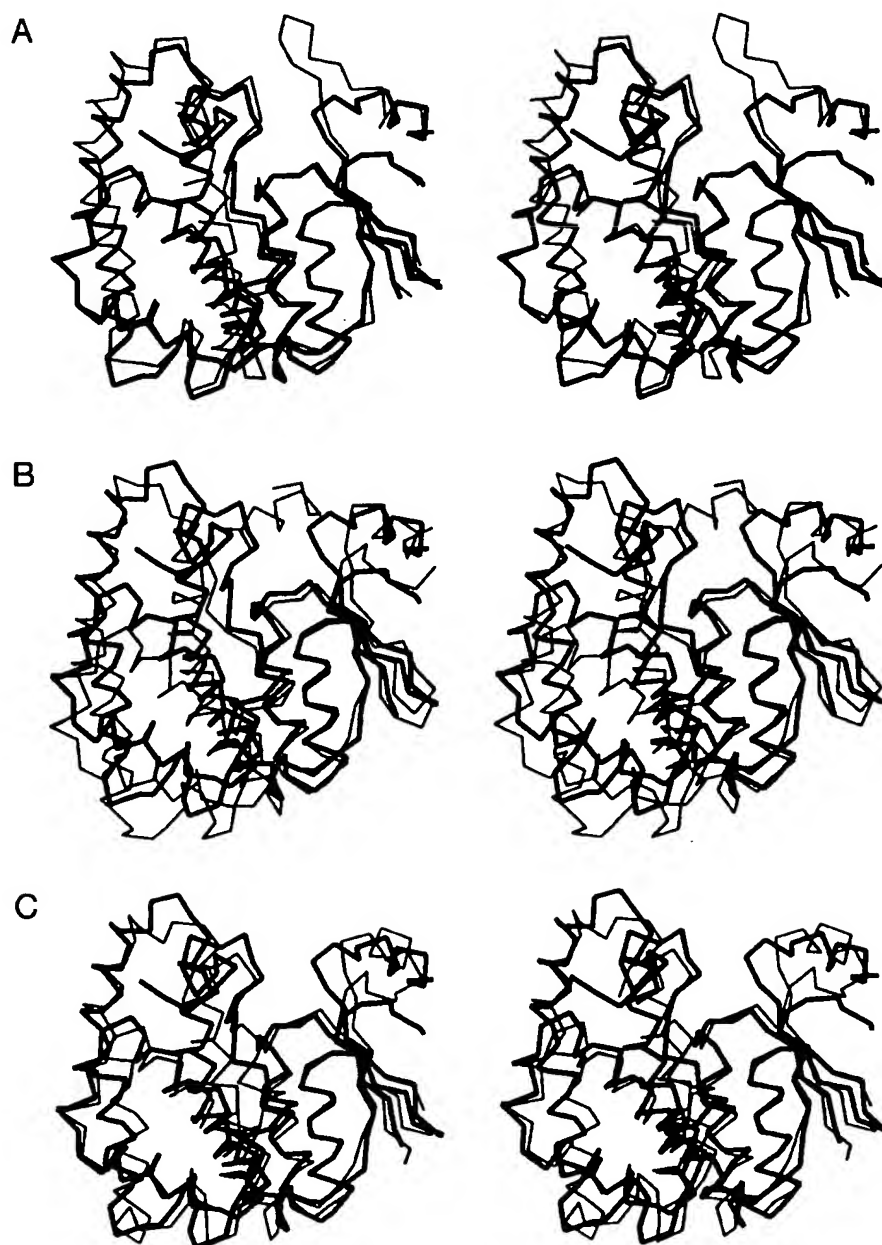


Fig. 9. Stereo views of $C\alpha$ atoms of *Sj* GST (in thick lines) with other GSTs (in thin lines) after the superposition of domain I. A: μ GST. B: α GST. C: π GST.

influenced by the presence of Trp or Phe residues. In *Sj* GST and μ GST, the carbonyl atom O_ω of glutathione cysteine forms a hydrogen bond with the indole ring atom $Ne1$ of Trp 7, and the backbone N atom of glutathione cysteine interacts with the Leu-Asn sequence. In contrast, the presence of Phe in α GST and π GST causes the peptide flip of glutathione cysteine, so that the carbonyl atom O_ω forms a hydrogen bond with the backbone atom N of Val 54 in α GST and Leu 50 in π GST. Furthermore, in comparison to the interaction at the γ -glutamate of glutathione, the amino acid residues that interact with the terminal carboxylate atoms of glutathione glycine are dissimilar in

their composition and orientation in the 4 GSTs. The Gln residue interacts with these carboxylate atoms in α GST and π GST in addition to the Arg, Lys, and Trp residues.

The orientation and alignment of glutathione thus enables the close association of the ring OH atom of residue Tyr 6 with the S_γ atom of glutathione in all 4 GSTs. A notable conformational difference of this residue between *Sj* GST and μ GST after the superposition of the binding site is the angle of ring planes. Relative to *Sj* GST, this angle is -34° (+ is counterclockwise, looking away from the backbone atoms); the angle is -15° for α GST and -10° for π GST. It has been suggested that in μ GST,

Table 2. Superposition of domain 2 after the alignment of domain 1 for the 4 GST structures^a: RMS deviation, rotation angle, and magnitude of translation

Fixed GST	Rotating GST		
	μ	α	π
<i>Sj</i>	1.66 Å, 5.5°, 1.10 Å	1.43 Å, 6.8°, 3.45 Å	1.42 Å, 5.1°, 1.27 Å
μ		1.91 Å, 7.2°, 4.52 Å	1.25 Å, 4.9°, 1.32 Å
α			1.64 Å, 8.5°, 3.98 Å

^a With the amino acid sequences aligned, the C α atoms of domain 1 (residues 2–6, 14–22, 27–32, 56–59, 62–65, and 67–77 of *Sj* GST) were all superimposed, i.e., the monomer structures (subunit A) of μ , α , and π GSTs were rotated and translated onto the *Sj* GST structure with respect to these aligned residues. Then the C α atoms of α -helices of domain 2 (residues 85–110, 113–125, 129–136, 149–164, 173–184, 186–192 of *Sj* GST) were superimposed between different pair combinations of GSTs.

the side chain of Thr 13 that sits on the Tyr 6 ring enhances the hydrogen bonding effect and the p*K*_a of the glutathione cysteine (Liu et al., 1993). Valine is the corresponding residue in *Sj* GST, methionine in α GST, and cysteine in π GST.

Site-directed mutagenesis studies have revealed the importance of specific residues in the glutathione binding site. Manoharan and coworkers (1992) showed that the substitutions of binding site residues Arg 13, Gln 62, and Asp 96 in π GST resulted in 20–50-fold decreases of both the catalysis and glutathione binding in comparison to the wild type. However, the substitution of Tyr 7 to phenylalanine still resulted in 27% of the wild-type capacity to bind glutathione, but the enzymatic catalysis was reduced to less than 1%. The substantial reduction in the catalytic activity due to the substitution of tyrosine was also demonstrated in μ GST (Liu et al., 1992). Furthermore, there are differing views on how the thiol anion of glutathione is created to promote high nucleophilic reactivity, i.e., either the thiol of glutathione is deprotonated (Liu et al., 1992) or tyrosine OH group is deprotonated (Karshikoff et al., 1993). Moreover, the importance of binding of carboxylate anions of glutathione was also shown by the mutation substitution at Trp 38 in π GST, which decreased the glutathione binding activity (Nishihira et al. 1992). Nishihira and coworkers (1993) also showed that the hydrophobic ligand binding that may occur away from the glutathione binding site caused the quenching of Trp 38 fluorescence, which suggested a conformational change of the enzyme.

Although *Sj* GST is from an invertebrate species and μ GST from a mammal, the 2 proteins show remarkable similarities in the amino acid sequence and structure. The shared characteristics in the 3-dimensional structures of *Sj* GST and μ GST extend to the similarity in biochemical properties. Walker and coworkers (1993) compared the characteristics of *Sj* GST (26 kDa) and *Schistosoma mansoni* (*Sm*) GST (28 kDa); whereas the latter showed catalytic properties similar to the mammalian μ class GSTs, the former gave hybridized enzymatic properties of μ and α GSTs.

In the GST binding site, glutathione binds to the domain 1 side, and there is more space on the domain 2 side for additional binding by an electrophilic substrate. This region may be called hydrophobic (H-) subsite (Reinemer et al., 1991; Garcia-Sáez et al., 1994) because the residues in this region are generally hydrophobic, but notable differences are evident. For example, in μ GST the residues Val 9, Leu 12, Ile 111, Tyr 115, Phe 208, and Ser 209 were shown to be important for the binding of

1-(*S*-glutathionyl)-2,4,6-trinitrocyclohexadienyl (GTD) (Ji et al., 1993) and the diastereomers of 9-(*S*-glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene (Ji et al., 1994). The corresponding residues in *Sj* GST are Ile 9, Leu 12, Ser 106, Tyr 110, Gln 203, and Gly 204 (Kinemage 1). Here the key difference is the change of Phe 208 in μ GST to Gln 203 in *Sj* GST, for the latter is observed to interact with Arg 102. The side chain of Arg 102 point toward the binding site and glutathione. The side chain of corresponding residue Arg 107 in μ GST points away from the binding site and interacts with Asp 161 that is further away. The combination of Arg 102 and Gln 203 in *Sj* GST could make the binding of such a molecule as GTD even more specific in *Sj* GST.

In the α GST structure, the ring portion of the substrate *S*-benzyl-glutathione is surrounded by hydrophobic residues (Val 11, Leu 106, Leu 107, Ala 215, Phe 219, and Phe 221). The binding site of α GST is less open than that of the other GSTs due to its C-terminal α -helix. In π GST, Phe 8, Pro 9, Val 10, Ala 102, and Tyr 106 may be the possible residues for xenobiotic binding interaction. The recent structure of mouse π GST (Garcia-Sáez et al., 1994) complexed with *S*-nitrobenzyl-glutathione showed the specificity of these and other residues. Furthermore, Tyr 115 in μ GST is shown to be an important residue for catalytic activity (Johnson et al., 1993). The corresponding residue in *Sj* GST is Tyr 110, Tyr 105 in π GST, and Val 110 in α GST. The influence of this residue on catalysis in GSTs other than μ GST has yet to be determined.

Fusion peptide and antibody reactivity to *Sj* GST

The fusion peptide of gp41 forms a series of tight turns (Fig. 8; Kinemage 1). Muster and coworkers (1993) showed that this peptide fused to *Sj* GST reacted with the Mab 2F5, which strongly suggests a recognizable conformation in solution. Compared to a free peptide, the fusion peptide may be stabilized by favorable interactions with the surface of the GST protein. Computer modeling predicts that the sequence Glu-Leu-Asp-Lys-Trp-Ala is a part of an α -helix (Gallaher et al., 1989). Because only this short sequence was utilized in the fusion with *Sj* GST, the predicted structure may not be expected in the present crystal lattice. Moreover, a recent structure of gp120 peptide in an antibody-antigen complex also adopts a conformation of tight turns (Ghiara et al., 1994). Once determined, the larger context of

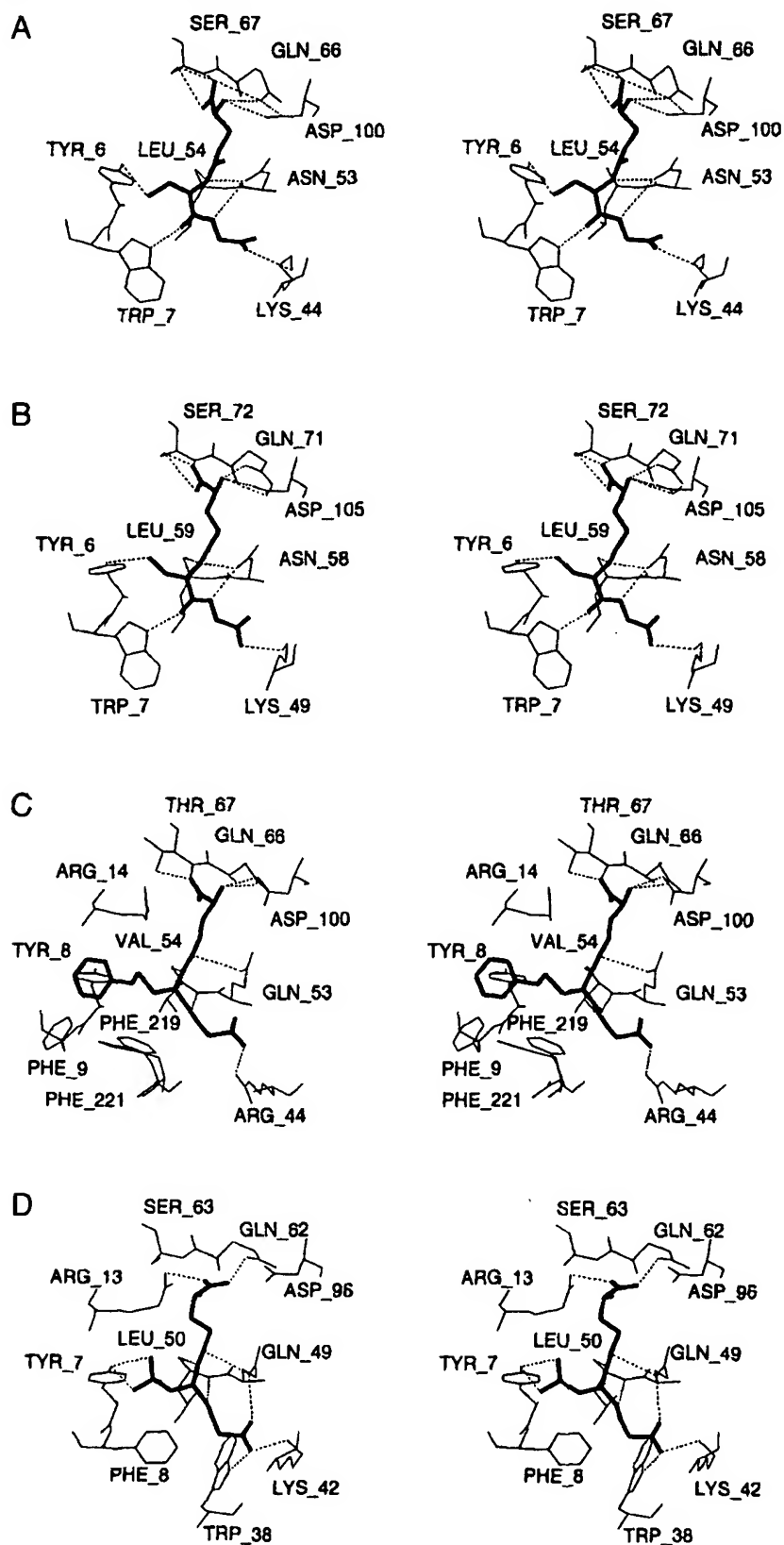


Fig. 10. Glutathione substrate (in thick lines) in the GST binding site. A: Glutathione in *Sj* GST. B: Glutathione in μ GST. C: S-benzyl-glutathione in α GST. D: Glutathione sulfonate in π GST.

HIV-1 protein structures should help to clarify the conformational significance of these peptides.

The *Sj* GST has been identified as a specific target for the host immune response in resistant mice (Smith et al., 1986). Although the anti-*Sj* GST antibody reactivity in infected humans has not been conclusive (Davern et al., 1990), more work can be done to determine the neutralizing characteristics of the antibodies against *Sj* GST and other components in schistosomiasis. For example, one of the monoclonal antibodies employed (250-7) by Davern and coworkers (1990) has an epitope specific for the C-terminal residues of *Sj* GST (sequence GGGDHPPK). In contrast, the binding by this antibody was weaker for the 26-kDa *Sm* GST with the C-terminal residues GGGDAPPK. The change of amino acid from histidine in *Sj* GST to alanine in *Sm* GST presumably has a significant effect on the epitope recognized by the antibody. The location of the C-terminal residues of the *Sj* GST structure presented here is at the surface of the protein; thus, these residues are accessible as epitopes for antibody recognition. The binding specificity conferred by His 214 (and perhaps also by Asp 213) may indicate that the antibody-antigen interaction could be electrostatic in nature. Because the inhibition of these and other related GSTs would offer the possibility of combining chemotherapy and immunotherapy (Mitchell, 1989; Trottein et al., 1992; Capron & Capron, 1994), additional 3-dimensional structural studies of the proteins from parasitic helminths would further enhance the development of a vaccine against the transmission of schistosomiasis.

Materials and methods

Expression and purification of *Sj* GST

The plasmid pGEX-2T was used to express the *Sj* GST fused with the gp41 peptide (Smith et al., 1986, 1987; Smith & Johnson, 1988; Muster et al., 1993; Pharmacia LKB Biotechnology, Uppsala, Sweden) in the *Escherichia coli* strain MC10611. In this vector, the gene coding for GST is under transcriptional control of the *tac* promoter. Isopropyl- β -D-thiogalactoside (IPTG) was added at a concentration of 1 mM for induction of the *tac* promoter. The bacterial cells were lysed by sonication and the lysate was cleared by centrifugation. The cleared lysate was applied to a 2-mL glutathione-Sepharose B column (Pharmacia LKB) that was equilibrated with a buffer of 20 mM sodium phosphate, 150 mM NaCl, and 1% Triton X-100, pH 7.3. Bound *Sj* GST fusion protein was subsequently eluted with 20 mM glutathione (reduced), 100 mM Tris, 120 mM NaCl, pH 8. The elution fractions were analyzed by SDS-PAGE. Positive fractions were pooled. Typical yield was 10–15 mg of purified *Sj* GST fusion protein per liter of bacterial culture.

Crystallization and data collection

Crystallization of *Sj* GST fused with the gp41 peptide was performed with hanging-drop vapor diffusion method. A 5- μ L volume of protein solution (10 mg/mL) was combined with 5 μ L of precipitating stock solution of 60% (w/v) PEG 3350 (polyethylene glycol) and 50 mM potassium phosphate, pH 6.5. This droplet was equilibrated to a 1-mL reservoir of the same PEG solution at 22 °C. Long, needle-like crystals appeared within 2 weeks.

The crystals have Laue symmetry of $4/mmm$ and belong to the space group $P4_32_12$ with cell dimensions $a = b = 94.7$ Å, and $c = 58.1$ Å. There is 1 GST monomer per asymmetric unit. X-ray diffraction data were collected with Rigaku RAXIS-II Imaging Plate Diffractometer mounted on Rigaku RU200 X-ray Generator. The crystal to detector distance was set at 133.4 mm and generator settings were 40 kV and 70 mA. The data set contained 8,485 unique reflections with $I > 1\sigma(I)$ to 2.50 Å (86.9% complete), and 10,057 reflections with $I > 1\sigma(I)$ to 2.25 Å (75.9% complete). Reflections in the 2.50–2.25 Å range were recorded at the corners of the imaging plates and consequently only 45.1% complete. The R_{merge} value was 11.7% for all observations with $I > 1\sigma(I)$.

Molecular replacement

The structure of *Sj* GST was determined by molecular replacement (Rossmann, 1990) using the unmodified coordinates of the subunit A of the previously determined μ GST (Ji et al., 1992) as the starting model. The program package MERLOT (Fitzgerald, 1988) was used in these calculations. The rotation function gave 2 different solutions at different resolution limits. Both of these 2 orientations were used to calculate translation functions, but only one gave consistent peaks at various Harker sections. At the same time, the space group ambiguity was resolved. This solution was also tested by program RTMAP (written by J.X. Ho, unpubl.) that maps crystallographic R factors and correlation coefficients over all possible translation vectors for a given range of orientation angles. The translation vector that corresponded to the best R factor (47.9%) and highest correlation coefficient (30.2%) was consistent with the results from Patterson translation function for 2,248 reflections with $F_o > 3\sigma(F_o)$ in the resolution range of 20.0–4.0 Å. This model was then used in a rigid body refinement after which the R factor was 30.4% and correlation coefficient 48.6% for the same group of data. The crystal packing of the molecules produces a perfect dimer of GST, i.e., 2 monomers relate to each other by a crystallographic 2-fold rotation symmetry.

Structure refinement

After the rotation and translation search, the secondary structure elements of the GST model were adjusted in the rigid body refinement performed with XPLOR program (Brünger, 1991). Then initial side-chain changes were made after alignments of the amino acid sequences of *Sj* GST and μ GST (Fig. 3). Direct changes were made for side chains similar in size and polarity such as Ile and Leu, and Met and Lys; other residues with pairings that did not share such characteristics were converted into alanines to produce the initial *Sj* GST model. Next, a simulated annealing calculation was performed with XPLOR using reflections in the 8.0–3.0-Å resolution range. Omit maps were calculated for appropriate replacements of side chains and adjustments of their geometries. Molecular graphics programs SIDUS (Lim, 1994) and O (Jones et al., 1991) were used to make these changes. Later, the resolution range was altered to include the 6–2.3-Å data. As the refinement progressed, water oxygen atoms were assigned for electron density peaks greater than 3σ on the $F_o - F_c$ contour maps. The present *Sj* GST structure should be defined as one determined to 2.5 Å resolution; although the diffraction data extended beyond 2.5 Å resolution,

less than 50% of the data were present in the higher resolution range. The final refinement of the *Sj* GST structure was performed with the restrained least-squares refinement program GPRLSA (Hendrickson & Konnert, 1980; Furey et al., 1982; Hendrickson, 1985). To determine the structure of fused gp41 peptide, the main structure of *Sj* GST was held fixed and the C-terminal residues were allowed to move in the simulated annealing calculations.

Determination of the main *Sj* GST structure

The *Sj* GST and μ GST proteins share 42% amino acid pair identities and many other homologous pairs (Fig. 3). One notable difference observed from the amino acid alignment is that *Sj* GST does not have a long loop that is present in μ GST (residues 33–42) in domain 1. Initial rigid body refinements performed without this loop resulted in the translation of as much as 1 Å for some of the α -helices, indicating that the tertiary structures of the 2 domains of *Sj* GST were different from those of μ GST. Furthermore, the 78–84 loop connecting the α -helices 3 and 4 was modified a few times before it was stabilized and showed a different spatial placement with respect to the 2 neighboring α -helices than the corresponding 83–89 loop in μ GST. Subtle structural differences were also noted, as when residue 154 was not changed from isoleucine of μ GST into alanine by oversight during the first several rounds of refinement. Omit map for this residue position did not fit well for isoleucine.

During the initial stages of refinement, glutathione was not included in the model. However, $F_o - F_c$ electron density map clearly indicated its presence in the binding site (Fig. 5A). Glutathione was added, along with additional water molecules where appropriate. The presence of oxidized glutathione (2 glutathiones joined by a disulfide bond) was ruled out. As higher resolution data were included in the refinement, i.e., reflections in the 6.0–2.3-Å range, additional water molecules were gradually located.

Difficulty in the *Sj* GST structure refinement was encountered with the positioning of the C-terminal residues. The sequence alignment indicated that the 3 glycine residues 210–212 may be paired with the last 3 residues of μ GST (215–217), but this left the additional 5 residues of *Sj* GST (His 213, Asp 214, Pro 215, Pro 216, and Lys 217) and another 15 residues of the fusion peptide without a starting model. This question was resolved by the refinement performed (using reflections in the 8.0–3.0-Å range) without the residues beyond the α -helix 8. The C-terminus of *Sj* GST was gradually rebuilt and showed the same loop folding pattern as that for the μ GST. The presence of electron densities for 3 glycines (210–212) supported for this folding pathway. The $F_o - F_c$ density in *Sj* GST could be traced to residue Pro 215 without addition of water molecules. A similar difficulty was also encountered for the structure of μ GST (Ji et al., 1992), for which the interpretation of the last 17 residues was resolved when 150 water molecules were added.

Determination of the fusion peptide structure

After the *Sj* GST structure was determined to residue Pro 215, the refinement was extended to higher resolution reflections in the 6.0–2.3-Å range. With the addition of approximately 50 water oxygen atoms, the location of the fusion peptide structure became more evident. Its electron density was located in the

pocket formed by the α -helices 6, 7, and 8, and near Pro 215, but it was not connected to Pro 215. Examination of the crystal symmetry indicated 2 possibilities: (1) the fusion peptide extends from Pro 215 directly to this density so that it lies on the surface of the main *Sj* GST structure, or (2) it extends to the density in the neighboring region so that the last several residues of the fusion peptide would interact with symmetry-related GSTs other than the original *Sj* GST in the main cell (Fig. 7). The second possibility was a more appropriate choice, because when an alanine model was placed in the fusion peptide region, the extension of electron density was evident from the neighboring region to Pro 215. In contrast, the direct connection of the electron density on the surface of *Sj* GST to Pro 215 could not be made even after many rounds of model building and refinement. Following this determination, residues were added at the C-terminus to Arg 223. At this point in the sequence, the fusion peptide was now interacting with the symmetry-related GSTs. The $F_o - F_c$ map indicated that the last 9 or 10 residues of the fusion peptide formed a loop structure. The residue Gly 224 initiates a turn, and at Met 226 another sharp turn is made. Finally, there is another turn at Asp 229 and Lys 230 (Fig. 8).

Acknowledgments

This research was supported by a grant from the Microgravity Science and Application Division of the National Aeronautics and Space Administration. K.L. is a National Research Council postdoctoral fellow. J.X.H. and K.K. were supported under contract with the Universities Space Research Association and Science and Technology Corporation, respectively. We thank the computing center at Marshall Space Flight Center for generous computing times. F.R. thanks Annelies Klima for excellent technical support and the funding by the Austrian Fonds zur Foerderung der wissenschaftlichen Forschung, project P8868-MOB.

References

- Armstrong RN. 1991. Glutathione S-transferases: Reaction mechanism, structure, and function. *Chem Res Toxicol* 4:131–140.
- Brophy PM, Barrett J. 1990. Glutathione transferase in helminths. *Parasitology* 100:345–349.
- Brophy PM, Crowley P, Barrett J. 1990. Relative distribution of glutathione transferase, glyoxalase I and glyoxalase II in helminths. *Int J Parasitol* 20:259–261.
- Brünger AT. 1991. Simulated annealing in crystallography. *Annu Rev Phys Chem* 42:197–223.
- Capron M, Capron A. 1994. Immunoglobulin E and effector cells in schistosomiasis. *Science* 264:1876–1877.
- Capron A, Dessaint JP, Capron M, Ouma JH, Butterworth AE. 1987. Immunity to schistosomes: Progress toward vaccine. *Science* 238:1065–1072.
- Carter DC, Rüker F, Ho JX, Lim K, Keeling K, Gilliland GL, Ji X. 1994. Fusion proteins as alternate crystallization paths to difficult structure problems. *Protein Pept Lett* 1:175–178.
- Davern KM, Tiu WU, Samaras N, Gearing DP, Hall BE, Garcia EG, Mitchell GF. 1990. *Schistosoma japonicum*: Monoclonal antibodies to the M_r 26,000 schistosome glutathione S-transferase (Sj26) in an assay for circulating antigen in infected individuals. *Exp Parasitol* 70:293–304.
- Dominey RJ, Nimmo IA, Cronshaw AD, Hayes JD. 1991. The major glutathione S-transferase in salmonid fish livers is homologous to the mammalian pi-class GST. *Comp Biochem Physiol* 100 B:93–98.
- Fainsod A, Margalit Y, Haffner R, Gruenbaum Y. 1991. Non-immunological precipitation of protein-DNA complexes using glutathione-S-transferase fusion proteins. *Nucleic Acids Res* 19:4005.
- Fitzgerald PM. 1988. MERLOT, an integrated package of computer programs for the determination of crystal structures by molecular replacement. *J Appl Crystallogr* 21:273–278.
- Furey W, Wang BC, Sax M. 1982. Crystallographic computing on an array processor. *J Appl Crystallogr* 15:160–166.
- Gallagher WR, Ball JM, Garry RF, Griffin MC, Montelaro RC. 1989. A general model for the transmembrane proteins of HIV and other retroviruses. *AIDS Res Hum Retroviruses* 5:431–440.
- García-Sáez I, Párraga A, Phillips MF, Mantle TJ, Coll M. 1994. Molecu-

- lar structure at 1.8 Å of mouse liver class pi glutathione S-transferase complexed with S-(π -nitrobenzyl)glutathione and other inhibitors. *J Mol Biol* 237:298–314.
- Ghiara JB, Stura EA, Stanfield RL, Proffy AT, Wilson IA. 1994. Crystal structure of the principal neutralization site of HIV-1. *Science* 264:82–85.
- Gilliland GL. 1993. Glutathione proteins. *Curr Opin Struct Biol* 3:875–884.
- Hendrickson WA. 1985. Stereochemically restrained refinement of macromolecular structures. *Methods Enzymol* 115:252–270.
- Hendrickson WA, Konnert JH. 1980. Stereochemically restrained crystallographic least-squares refinement of macromolecule structures. In: Srinivasan R, ed. *Biomolecular structure, conformation, function and evolution*, vol 1. Oxford: Pergamon Press. pp 43–57.
- Hiratsuka A, Sebata N, Kawashima K, Okuda H, Ogura K, Watabe T, Satoh K, Hatayama I, Tsuchida S, Ishikawa T, Sato K. 1990. A new class of rat glutathione S-transferase Yrs-Yrs inactivating reactive sulfate esters as metabolites of carcinogenic arylmethanols. *J Biol Chem* 265:11973–11981.
- Hughes AL. 1993. Rates of amino acid evolution in the 26- and 28-kDa glutathione S-transferases of *Schistosoma*. *Mol Biochem Parasitol* 58:43–52.
- Ji X, Armstrong RN, Gilliland GL. 1993. Snapshots along the reaction coordinate of an S_NAr reaction catalyzed by glutathione transferase. *Biochemistry* 32:12949–12954.
- Ji X, Johnson WW, Muctarr AS, Dickert L, Prasad SM, Ammon HL, Armstrong RN, Gilliland GL. 1994. Structure and function of the xenobiotic substrate binding site of a glutathione S-transferase as revealed by X-ray crystallographic analysis of product complexes with the diastereomers of 9-(S-glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene. *Biochemistry* 33:1043–1052.
- Ji X, Zhang P, Armstrong RN, Gilliland GL. 1992. The three-dimensional structure of a glutathione S-transferase from the mu gene class. Structural analysis of the binary complex of isoenzyme 3-3 and glutathione at 2.2 Å resolution. *Biochemistry* 31:10169–10184.
- Johnson WW, Liu S, Ji X, Gilliland GL, Armstrong RN. 1993. Tyrosine 115 participates both in chemical and physical steps of the catalytic mechanism of a glutathione S-transferase. *J Biol Chem* 268:11508–11511.
- Jones TA, Zou JY, Cowan, Kjeldgaard M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47:110–119.
- Karshikoff A, Reinemer P, Huber R, Ladenstein R. 1993. Electrostatic evidence for the activation of the glutathione thiol by Tyr 7 in π -class glutathione transferases. *Eur J Biochem* 215:663–670.
- Kong KH, Inoue H, Takahashi K. 1993. Site-directed mutagenesis study on the roles of evolutionally conserved aspartic acid residues in human glutathione S-transferase P1-1. *Protein Eng* 6:93–99.
- Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 24:946–950.
- Lim K. 1994. SIDUS: A new software for simplified approach to crystallographic and molecular modeling studies. *Thirteenth Annual Conference of the Molecular Graphics Society*, July 9–13, 1994, Evanston, Illinois.
- Liu S, Ji X, Gilliland GL, Stevens WJ, Armstrong RN. 1993. Second-sphere electrostatic effects in the active site of glutathione S-transferase. *J Am Chem Soc* 115:7910–7911.
- Liu S, Zhang P, Ji X, Johnson WW, Gilliland GL, Armstrong RN. 1992. Contribution of tyrosine 6 to the catalytic mechanism of isoenzyme 3-3 of glutathione S-transferase. *J Biol Chem* 267:4296–4299.
- Mannervik B, Awasthi YC, Board PG, Hayes JD, Di Ilio C, Ketterer B, Listowsky I, Morgenstern R, Muramatsu M, Pearson WR, Pickett CB, Sato K, Widerstern M, Wolf CD. 1992. Nomenclature for human glutathione transferases. *Biochem J* 282:305–308.
- Mannervik B, Danielson UH. 1988. Glutathione transferases – Structure and catalytic activity. *CRC Crit Rev Biochem* 23:283–337.
- Manoharan TH, Gulick AM, Reinemer P, Dirr HW, Huber R, Fahl WE. 1992. Mutational substitution of residues implicated by crystal structure in binding the substrate glutathione to human glutathione S-transferase π . *J Mol Biol* 226:319–322.
- Menéndez-Arias L, Young M, Oroszlan S. 1992. Purification and characterization of the mouse mammary tumor virus protease expressed in *Escherichia coli*. *J Biol Chem* 267:24134–24139.
- Meyer DJ, Coles B, Pemble SE, Gilmore KS, Fraser GM, Ketterer B. 1991. Theta, a new class of glutathione transferases purified from rat and man. *Biochem J* 274:409–414.
- Mignogna G, Allocati N, Aceto A, Piccolomini R, Di Ilio C, Barra D, Martini F. 1993. The amino acid sequence of glutathione transferase from *Proteus mirabilis*, a prototype of a new class of enzymes. *Eur J Biochem* 211:421–425.
- Mitchell GF. 1989. Glutathione S-transferases – Potential components of anti-schistosome vaccines? *Parasitol Today* 5:34–37.
- Muster T, Steindl F, Purtscher M, Trkola A, Klima A, Himmler G, Rüker F, Katinger H. A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1. 1993. *J Virol* 67:6642–6647.
- Nishihira J, Ishibashi T, Sakai M, Nishi S, Kumazaki T. 1992. Evidence for the involvement of tryptophan 38 in the active site of glutathione S-transferase P. *Biochem Biophys Res Commun* 185:1069–1077.
- Nishihira J, Ishibashi T, Sakai M, Tsuda S, Hikichi K. 1993. Identification of the hydrophobic ligand-binding region in recombinant glutathione S-transferase P and its binding effect on the conformational state of the enzyme. *Arch Biochem Biophys* 302:128–133.
- Oshima-Hirayama N, Yoshikawa K, Nishioka T, Oda J. 1993. Lipase from *Pseudomonas aeruginosa*: Production in *Escherichia coli* and activation in vitro with a protein from the downstream gene. *Eur J Biochem* 215:239–246.
- Reinemer P, Dirr HW, Ladenstein R, Huber R, Lo Bello M, Federici G, Parker MW. 1992. Three-dimensional structure of class π glutathione S-transferase from human placenta in complex with S-hexylglutathione at 2.8 Å resolution. *J Mol Biol* 227:214–226.
- Reinemer P, Dirr HW, Ladenstein R, Schäffer J, Gally O, Huber R. 1991. The three-dimensional structure of class π glutathione S-transferase in complex with glutathione sulfonate at 2.3 Å resolution. *EMBO J* 10:1997–2005.
- Rossmann MG. 1990. The molecular replacement method. *Acta Crystallogr A* 46:73–82.
- Rushmore TH, Pickett CB. 1993. Glutathione S-transferases, structure, regulation, and therapeutic implications. *J Biol Chem* 268:11475–11478.
- Sher A, James SL, Correa-Oliveira R, Hieny S, Pearce E. 1989. Schistosome vaccines: Current progress and future prospects. *Parasitology* 98:S61–S68.
- Sinning I, Kleywegt GJ, Cowan SW, Reinemer P, Dirr HW, Huber R, Gilliland GL, Armstrong RN, Ji X, Board PG, Olin B, Mannervik B, Jones TA. 1993. Structure determination and refinement of human alpha class glutathione transferase A1-1, and a comparison with the mu and pi class enzymes. *J Mol Biol* 232:192–212.
- Slany RK, Bösl M, Crain PF, Kersten H. 1993. A new function of S-adenosylmethionine: The ribosyl moiety of AdoMet is the precursor of the cyclopentenadiol moiety of the tRNA wobble base queuine. *Biochemistry* 32:7811–7817.
- Smith DB, Davern KM, Board PG, Tiu WU, Garcia EG, Mitchell GF. 1986. M_r 26,000 antigen of *Schistosoma japonicum* recognized by resistant WEHI 129/J mice is a parasite glutathione S-transferase. *Proc Natl Acad Sci USA* 83:8703–8707.
- Smith DB, Davern KM, Board PG, Tiu WU, Garcia EG, Mitchell GF. 1987. Correction. *Proc Natl Acad Sci USA* 84:6541.
- Smith DB, Johnson KS. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 67:31–40.
- Trottein F, Godin C, Pierce RJ, Sellin B, Taylor MG, Gorillot I, Silva MS, Lecocq JP, Capron A. 1992. Inter-species variation of schistosome 28-kDa glutathione S-transferases. *Mol Biochem Parasitol* 54:63–72.
- Tsuchida S, Sato K. 1992. Glutathione transferases and cancer. *CRC Rev Biochem Mol Biol* 27:337–384.
- Walker J, Crowley P, Moreman AD, Barrett J. 1993. Biochemical properties of glutathione S-transferases from *Schistosoma mansoni* and *Schistosoma japonicum*. *Mol Biochem Parasitol* 61:255–264.